PATENT

Amendments to the Specification:

Please replace paragraph [0001] of the specification with the following rewritten paragraph:

[0001] This application is related to U.S. Patent Application Nos. 09/574,268, entitled "Assist Channel Coding With Character Classifications" ~~(Attorney Docket No. D/AO038)~~, now issued as US 6,862,113; U.S. Patent Application No. 09/574,270, now issued as US 6,768,560, entitled "Assist Channel Coding With Vertical Block Error Correction" ~~(Attorney Docket No. D/AO039),~~; U.S. Patent Application No. 09/574,274,, entitled "Assist Channel Coding With Convolution Coding" ~~(Attorney Docket No. D/AO040),~~ now issued as US 6,628,837, and U.S. Patent Application No. 09/574,406, entitled "Assist Channel Coding Using A Rewrite Model" (Attorney Docket No. D/AO041), which are hereby incorporated by reference.

Please add paragraph [0015.1] as follows:

[0015.1] The functional utility of plain paper and other types of hardcopy documents can be enhanced significantly if the human readable information that they normally convey is supplemented by adding appropriate machine readable digital data to them. Input scanners can be employed for recovering this machine readable data, so the data can be employed for various purposes during the electronic processing of such documents and their human readable contents by electronic document processing systems, such as electronic copiers, text and graphic image processing systems, facsimile systems, electronic mail systems, electronic file systems, and document and character recognition equipment. Machine readable digital data can be recorded by writing two dimensional marks on a recording medium in accordance with a pattern which encodes the data either by the presence or absence of marks at a sequence of spatial locations or by the presence or absence of mark related transitions at such locations. The bar-like codes which others have proposed for recording digital data on paper utilize that type of encoding. Another approach is to encode machine readable

digital data in the shapes of the marks or "glyphs" that are written on the recording medium.

Please replace paragraph [0019] of the specification with the following rewritten paragraph:

[0019] FIG. 3 illustrates the uploading of the hardcopy document 106, with data from a primary channel 124 and a multilingual encoding 126 rendered thereon, from the hardcopy domain to the electronic domain. As shown in FIG. 3, multilingual decoding module 304 is comprised of an image extraction module 310, a decoding/decompression module 312 and a decoder module 316. In accordance with one aspect of the invention, a user first scans hardcopy document 106 into the system using scanner 308. A multilingual decoding module 304 then receives bitmap image or coded data 306 from scanner 308. An image extraction module 310 in the multilingual decoding module 304 separates the multilingual channel 118 data from bitmap primary channel 314 data. After a user selects the desired translation for output, a decoding/decompression module 312 in the multilingual decoding module 304 decodes and decompresses the appropriate portion of the multilingual channel 118 and passes the data along decoded and decompressed multilingual channel 319 to decoder module 316. The decoder module 316 performs OCR on the bitmap primary channel 314 and passes the primary channel data 112 to terminal 100. In a preferred embodiment, multilingual channel data 118 and 126 is comprised of data to assist the conversion of primary channel data 112 and 124, respectively, from a first human-readable language into at least a second human-readable language. It is also envisioned that multilingual channel data 118 and 126 may additionally include information helpful in performing OCR of the primary channel data 112 and 124, respectively. A method for assisting in performing OCR of primary channel data through the use of an assist channel is described in commonly assigned, co-pending U.S. Patent Application entitled, Assist Channel Coding with Character Classifications (Application No. 09/574,268), the contents of which are expressly incorporated by reference. An

"assist channel" of a hardcopy document is a machine readable encoding of side information that aids an OCR application in decoding the contents of a primary channel. The "primary channel" of a hardcopy document includes the human readable information of document. The primary channel, which cannot be modified and is slightly error prone to OCR processing, carries most of the information content of the document. One use of the assist channel is to encode information that assists in the identification of failures of an OCR application in decoding the contents of a primary channel. The assist channel may be encoded by partitioning the primary set of symbol data into a plurality of groups. The plurality of groups classify symbol data according to how likely the symbol data will occur in the hardcopy document and how likely symbol data in the primary set of symbol data are to be confused during processing of a scanned representation of the symbol data. The secondary set of encoding data is developed by associating the symbol data with ones of the plurality of groups. Once accurately reconstructed using the decoder module 316, the primary channel data 112 can be displayed on display 104 as image data 114.

Please replace paragraph [0024] of the specification with the following rewritten paragraph:

[0024] In a preferred embodiment, the encoded multilingual information appears on the face of the hardcopy document as a compact, visually benign representation of the primary information. Glyph marks represent one example of a suitable format for the encoded multilingual data. See, United States patent application of Dan S. Bloomberg, which was filed Jul. 31, 1990 under Ser. No. 07/560,514, entitled "Self Clocking Glyph Shape Codes," issued as US Patent 6,076,738. As shown in FIG. 4, glyph marks are composed of elongated slash-like marks or "glyphs" 422 and 423 that are written on a generally regular rectangular lattice of centers on a suitable recording medium 424. Suitably, the glyphs 422 and 423 are printed by a printer (not shown) operating at 300 d.p.i. to 600 d.p.i. to write 4 pixel x 4 pixel to 7 pixel x 7 pixel representations of the glyphs 422 and 423 on regularly spaced centers that are distributed widthwise

and lengthwise of the recording medium 424 to produce the code pattern 421. The glyphs of these fine grain glyph code patterns are not easily resolved by the unaided human eye when the code patterns are viewed under standard lighting conditions and at normal reading distances, so the code pattern 421 typically has a generally uniform gray scale appearance. Alternatively, the glyph marks may be modulated in an area to form a glyph half tone image or glyphtone as disclosed in commonly assigned U.S. Patent Nos. 5,315,098 and 5,706,099 the contents of which are expressly incorporated by reference. Nevertheless, the glyph code is still capable of effectively communicating machine readable digital information. To carry out this function, the glyphs 422 and 423 usually are tilted to the left and right, at about +45° and -45° with respect to the longitudinal dimension of the recording medium 424 to encode binary "1's" and "0's", respectively, as shown at 425.

Please replace paragraph [0025] of the specification with the following rewritten paragraph:

[0025] In a preferred embodiment, the encoded multilingual data represents a code C that describes a set of editing operations that can be applied to the primary information to convert it from a first (presentation) language into a second language. Assume that for each page of text P in a presentation language (e.g., English), there is an accurate translation ATL in each of a plurality of languages L, each with its own glyph description. Also assume a plurality of processing routines RL (perhaps one for each language) that can be applied to P to produce a translation of P into language L. The quality of this translation RL(P) may be anywhere on the continuum from very good to very bad. In any event, it is assumed that code C describes a set of editing functions E necessary to convert RL(P) into ATL. In the case where RL(P) closely approximates ATL, C will describe very minor, if any, editing functions. When RL(P) is very bad, C will describe more significant editing functions to apply to RL(P), making it identical to ATL. In other words, we compute a C such that ATL=E(C, RL(P)). Assuming that E and RL exist in multilingual

encoding/decoding modules 110 and 304 with an OCR engine, we need merely transmit C in glyphs on the page containing the human-readable text P. Multilingual encoding/decoding modules 110 and 304 would then reconstruct ATL by ~~OCR'ing~~ applying an OCR operation to P, applying RL to the result, and then correcting according to instructions in C. A method for reading and decoding a channel is described in commonly assigned, copending U.S. Patent Application entitled, ~~Multilingual encoding/decoding~~ "Assist Channel Coding with Character Classifications" (Application No. 09/574,268), the contents of which are hereby expressly incorporated by reference.